Some basic facts and definitions

1 Equivalence relation, partial order

In this section A denotes a set.

Definition. We say that R is a relation on A if $R \subset A \times A$. Sometimes we use notation xRy to say that $(x, y) \in R$.

Definition. Let R be a relation on A. We say that R is

1. *reflexive* if for every $x \in A$ one has xRx,
2. *symmetric* if xRy implies yRx,
3. *antisymmetric* if xRy and yRx (simultaneously) imply $x = y$,
4. *transitive* if xRy and yRz (simultaneously) imply xRz.

Definition. Let R be a relation on A. We say that R is an equivalence relation, usually denoted by $\sim$, if R is reflexive, symmetric and transitive. We say that R is a partial ordering, usually denoted by $\leq$, if R is reflexive, antisymmetric and transitive.

Definition. Given an equivalence relation $\sim$ on A, a non-empty subset S of A is called an equivalence class if for every $x \in S$ and $y \in S$ one has $x \sim y$. It is well-known and easy to check that equivalence classes are disjoint and that A is the union of them. Given $x \in A$ we denote the equivalence class corresponding to it by $\bar{x}$, i.e.

$$\bar{x} = \{y \mid y \sim x\}.$$

Definition. Let $\leq$ be a partial ordering on A. We say that A is partially-ordered set. A subset S of A is called a chain if for every $x \in S$ and $y \in S$ one has either $x \leq y$ or $y \leq x$. A subset S of A is called bounded if there exists $z \in A$ such that for every $x \in S$ one has $x \leq z$. An element $x \in A$ is called maximal if there is no $z \in A$, $z \neq x$ such that $x \leq z$.

Lemma 1.1 (Zorn's Lemma) *Let A be a partially-ordered set such that every chain in A is bounded. Then A has a maximal element.*

2 Linear spaces

Let $\mathbb{K}$ denotes either $\mathbb{R}$ or $\mathbb{C}$.

Definition. A set X is called a linear space over $\mathbb{K}$ if two operations $+ : X \times X \to X$ and $\cdot : \mathbb{K} \times X \to X$ are defined and satisfy

1. for every $x \in X$, $y \in X$ one has $x + y = y + x$,
2. for every $x \in X$, $y \in X$, $z \in X$ one has $(x + y) + z = x + (y + z)$,
3. there exists $\theta \in X$ such that for every $x \in X$ one has $x + \theta = x$,
4. for every $x \in X$ there exists $\tilde{x}$ such that $x + \tilde{x} = \theta$;
5. for every $\alpha \in K$, $x \in X$, $y \in X$ one has $\alpha \cdot (x + y) = (\alpha \cdot y) + (\alpha \cdot x)$,
6. for every $\alpha \in K$, $\beta \in K$, $x \in X$, one has $(\alpha + \beta) \cdot x = (\alpha \cdot x) + (\beta \cdot x)$,
7. for every $\alpha \in K$, $\beta \in K$, $x \in X$, one has $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$,
8. for every $x \in X$ one has $1 \cdot x = x$,

Remarks. 1. If such θ (as in "3") exists then exists and unique. Usually we denote it by 0. Do not confuse between $0 \in \mathbb{K}$ and $0 \in X$.
2. If such $\tilde{x}$ (as in "4") exists then exists and unique. We denote it by $-x$.
3. Uniqueness of $-x$ allows us to define operation $- : X \times X \to X$ as $x - y = x + (-y)$.
4. If $\alpha \in \mathbb{K}$ and $x \in X$ we usually omit "$\cdot$" in $\alpha \cdot x$ and write just αx.

Exercises. 1. Prove remarks 1 and 2.
2. Show that for every $\alpha \in \mathbb{K}$ and $x \in X$ one has $0 \cdot x = \theta$ and $\alpha \cdot \theta = \theta$.

Definitions. Let X be a linear space. A non-empty subset Y of X is called a subspace if for every α, $\beta \in K$, x, $y \in Y$ one has $\alpha x + \beta y \in Y$. Let S be a non-empty subset of X. The span of S, denoted by span S is defined by

$$\text{span } S = \left\{ \sum_{i=1}^{n} \alpha_i x_i \mid n \in \mathbb{N}, \alpha_i \in \mathbb{K}, x_i \in S \text{ for every } i \leq n \right\}.$$

In fact span S is the smallest subspace of X containing S (i.e. span S is the subspace and for every subspace $Y \subset X$ if $Y \supset S$ then $Y \supset \text{span } S$).

Definitions. Let X be a linear space. A finite non-empty set $S = \{x_1, ..., x_n\} \subset X$ is called linearly independent if $\sum_{i=1}^n \alpha_i x_i = 0$ implies $\alpha_1 = ... = \alpha_n = 0$. In general an (infinite) subset S of X is called linearly independent if any finite subset of S is linearly independent. If X is spanned by linearly independent set S (i.e. span $S = X$) then S called an algebraic (or Hamel) basis of X.

Theorem 2.1 *Every linear space $X \neq \{0\}$ has an algebraic basis.*

Remark. This theorem follows from Zorn's Lemma.

Definition. Let X be a linear space and S be its basis. If S is finite then X is called finite-dimensional and the dimension of X is the cardinality of S (it is well-known and easy to show that any two bases of X has the same amount of elements). If S is not finite then we say that X is infinite dimensional and the dimension of X is infinity. We denote the dimension of X by dim X. If $X = \{0\}$ we define dim $X = 0$.

3 Convex sets, Minkowski sum of sets

In this section X denotes a linear space.

Definitions. A set $A \subset X$ is called convex if for every $x \in A$, $y \in A$ and $\alpha \in [0, 1]$ one has $\alpha x + (1 - \alpha)y \in A$. The convex hull of A, denoted by conv A, is the smallest (in sense of inclusion) convex set containing A, i.e.

$$\text{conv } A = \left\{ \sum_{i=1}^{n} \alpha_i x_i \mid n \in \mathbb{N}, \alpha_i \geq 0, x_i \in A \text{ for every } i \leq n, \text{ and } \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

or, equivalently, conv A is the intersection of all convex sets containing A.

Definitions. *(Minkowski sum and difference of sets)* Let $A \subset X$, $B \subset X$, $x \in X$, $\alpha \in \mathbb{K}$. We define

$$A + B = \{x + y \mid x \in A, y \in B\}, \quad x + A = \{x\} + A = \{x + y \mid y \in A\},$$
$$\alpha A = \{\alpha x \mid x \in A\}, \quad -A = (-1)A = \{-x \mid x \in A\},$$
$$A - B = A + (-B) = \{x - y \mid x \in A, y \in B\}.$$

Remark. Note that in general $A + A \neq 2A$, $A - A \neq 0$.

Exercise. Let A be a convex subset of X. Let $\alpha \geq 0$, $\beta \geq 0$. Show that $\alpha A + \beta A = (\alpha + \beta)A$.

4 Metric spaces

Definition. Let X be a set and let $\rho : X \times X \to [0, \infty)$ be a function satisfying

1. $\rho(x, y) = 0$ if and only if $x = y$,
2. $\rho(x, y) = \rho(y, x)$ for every $x,\ y \in X$,
3. *(triangle inequality)* $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for every $x,\ y,\ z \in X$.

Then ρ is called a metric and $X = (X, \rho)$ is called a metric space. If Y is a subspace of X then

$$\bar{\rho} = \rho_{Y \times Y}$$

is a metric on Y which is called induced (by ρ) metric.

Exercise. Show that the condition "3" in the definition above is equivalent to the condition

$$|\rho(x, y) - \rho(x, z)| \leq \rho(z, y) \quad \text{for every} \quad x, y, z \in X. \tag{4.1}$$

Remark. Inequality (4.1) we also call triangle inequality.

Definitions. Let $X = (X, \rho)$ be a metric space. Open ball with center at $x \in X$ and radius $r > 0$, denoted by $B(x, r)$, is defined by

$$B(x, r) = \{y \in X \mid \rho(x, y) < r\}.$$

The family of all open balls defines a topology on X which is called induced (by ρ) topology. Equivalently we can define this topology defining open sets as follows: A set $A \subset X$ is called open if for every $x \in A$ there exists $r > 0$ such that $B(x, r) \subset A$. As usual a set $A \subset X$ is called closed if a complement of A is open (recall that complement of A, denoted by A^c, is the set of all points of X which are not in A. The interior of a set A, denoted by Int A, is defined as the union of all open sets contained in A, i.e. $x \in \text{Int } A$ if and only if there exists $r > 0$ such that $B(x, r) \subset A$. The closure of A, denoted by clos A, is set of all point x such that for every $r > 0$ one has $A \cap B(x, r) \neq \emptyset$. The boundary of a set A, denoted by ∂A, is defined as set of all points which are in clos A but not in Int A, i.e. $\partial A = \text{clos } A \cap (\text{Int } A)^c$.

Definition. Let $X = (X, \rho)$ be a metric space and $\{x_i\}_{i=1}^{\infty}$ be a sequence in X. We say that the sequence converges to $x_0 \in X$ if for every $\varepsilon > 0$ there exists $N = N(\varepsilon)$ such that for every $n \geq N$ one has $\rho(x_n, x_0) < \varepsilon$. We write

$$\lim_{i \to \infty} x_i = x_0 \quad \text{or} \quad x_i \to x_0$$

and say that x_0 is the limit of the sequence. If a sequence converges to some limit then it is called convergent, otherwise it is called divergent.

Definition. Let $X = (X, \rho)$ be a metric space, $A \subset X$, $x_0 \in X$. We say that x_0 is a limit point of A if there exists a sequence $\{x_i\}_{i=1}^{\infty}$ in A such that $x_i \neq x_0$ for every i and $\lim_{i \to \infty} x_i = x_0$.

Claim 4.1 *Let $X = (X, \rho)$ be a metric space and $A \subset X$. Then A is closed if and only if $A = \text{clos } A$. Moreover, A is closed if and only if A contains all its limit points.*

Exercise. Let $x \in X$ and $r > 0$. Show that

$$\text{clos } B(x, r) = \{y \in X \mid \rho(x, y) \leq r\}.$$

Exercises. Let $\|\cdot\|$ be a norm on X and ρ be the induced metric (i.e. $\rho(x, y) = \|x - y\|$ for every x, y in X).
1. Show that $B(x, r)$ and clos $B(x, r)$ are convex for every $x \in X$ and $r > 0$.
2. Show that for every x, $y \in X$ every $r > 0$, $R > 0$, and every α, $\beta \in \mathbb{K}$ ($\beta \neq 0$) one has

$$\alpha B(x, r) + \beta B(y, R) = B\left(\alpha x + \beta y, |\alpha| r + |\beta| R\right),$$

$$\alpha \, \text{clos } B(x, r) + \beta B(y, R) = B\left(\alpha x + \beta y, |\alpha| r + |\beta| R\right),$$

$$\alpha \, \text{clos } B(x, r) + \beta \, \text{clos } B(y, R) = B\left(\alpha x + \beta y, |\alpha| r + |\beta| R\right),$$

where sums of sets and product of a set by a scalar are in the sense of definitions of Section 3.

5 Compactness in metric spaces

In this section $X = (X, \rho)$ denotes a metric space.

Definitions. Let K be a subset of X. K is called compact if every covering of K by open sets contains a finite subcovering, i.e. if

$$K \subset \bigcup_{U \in \mathcal{F}} U, \quad \text{every } U \in \mathcal{F} \text{ is an open subset of } X$$

implies

$$K \subset \bigcup_{i=1}^{n} U_i$$

for some U_1, ..., $U_n \in \mathcal{F}$.

Remark. Note that the compactness is a topological property.

Definition. Let K and A be subsets of X. Let $\varepsilon > 0$. A is called an ε-net for K if K can be covered by balls of radius ε with centers in A, i.e. if

$$K \subset \bigcup_{a \in A} B(a, \varepsilon).$$

If A is finite we say that there exists a finite ε-net for K.

Here are some basic facts about compact sets.

Lemma 5.1 *Let $A \subset X$ be a closed set and $K \subset X$ be a compact set. If $A \subset K$ then A is compact.*

Lemma 5.2 *Let $K \subset X$ be a compact set and $x \in X$. If $x \notin K$ then there exist open disjoint sets U and V such that $x \in U$ and $K \subset V$.*

Lemma 5.3 *Let $K \subset X$ be a compact set. Then K is closed.*

Definition. Let $A \subset X$. A is called sequentially compact if every sequence in A has a subsequence that converges to a point in A.

Theorem 5.4 *Let $K \subset X$. The following are equivalent*

[i] K is compact.

[ii] K is sequentially compact.

[iii] K is complete (as a space (K, ρ)) and for every $\varepsilon > 0$ there exists a finite ε-net for K.

Theorem 5.5 *Let X_1, ..., X_n be topological spaces (as usual it is enough for us to consider metric spaces only). Let $K_i \subset X_i$, $i \le n$, be compact sets. Then $K_1 \times K_2 \times ... \times K_n$ is a compact set in $X_1 \times X_2 \times ... \times X_n$. (Topology on $X_1 \times X_2 \times ... \times X_n$ is defined by $\{U_1 \times U_2 \times ... \times U_n\}$ over all choices of open sets $U_i \subset X_i$, $i \le n$. Equivalently, a sequence*

$$f^m = (f_1^m, f_2^m, ..., f_n^m) \to f = (f_1, f_2, ..., f_n)$$

if and only if $\lim_{m\to\infty} f_i^m = f_i$ for every $i \le n$).

At the end of this section we discuss relatively compact sets.

Definition. A subset K of X is called relative compact if clos K is compact.

Example. Let $X = \mathbb{R}$ with the standard metric (i.e. $\rho(x, y) = |x-y|$). Then a segment $[0, 1]$ is compact subset of X; a segment $[0, 1)$ is not compact, but relatively compact.

Remark. Note that Lemma 5.3 implies that every compact set is also relatively compact.

By Lemma 5.1 we immediately obtain

Lemma 5.6 *Let A be a subset of X and $K \subset X$ be a relatively compact set. If $A \subset K$ then A is relative compact.*

Theorem 5.5 has the following two corollaries.

Corollary 5.7 *Every relatively compact set is bounded (i.e. is contained in some ball).*

Corollary 5.8 *Let X be complete metric space and $K \subset X$. The following are equivalent*

[i] K is a relatively compact set.

[ii] For every $\varepsilon > 0$ there exists a finite ε-net for K.

[iii] For every $\varepsilon > 0$ there exists a compact ε-net for K.

Exercises. Prove Lemma 5.6 and Corollaries 5.7 and 5.8

6 Continuous functions on metric spaces

In this section $X = (X, \rho)$, $Y = (Y, \sigma)$ denote metric spaces and $f : X \to Y$ denotes a function. In most cases we will work with $Y = \mathbb{K}$ (and $\sigma(x, y) = |x - y|$).

Definitions. 1. A function f is continuous at the point $x \in X$ if for every sequence $x_i \to x$ one has $f(x_i) \to f(x)$.
2. A function f is continuous at the point $x \in X$ if for every $\varepsilon > 0$ there exists $\delta = \delta(x, \varepsilon) > 0$ such that for every y satisfying $\rho(y, x) < \delta$ (i.e. for every $y \in B(x, \delta)$) one has $\sigma(f(y), f(x)) \leq \varepsilon$ (i.e. $f(y) \in B(f(x), \varepsilon)$).

Exercise. Show that definitions 1 and 2 are equivalent.

Definitions. 3. A function f is continuous if it is continuous at every point.
4. A function f is continuous if for every open set $U \subset Y$ its preimage $f^{-1}(U)$ is open in X.
5. A function f is continuous if for every closed set $V \subset Y$ its preimage $f^{-1}(V)$ is closed in X.

Exercise. Show that definitions 3, 4, and 5 are equivalent.

Theorem 6.1 *Let A be a relative compact in X and K be a compact in X. Let $f : X \to Y$ be a continuous function. Then $f(A)$ is a relative compact in Y and $f(K)$ is a compact in Y.*

Theorem 6.2 *Let K be a compact in X and $f : X \to \mathbb{K}$ be a continuous function. Then f attains its minimal and maximal value on K, that is there exist x, $y \in K$ such that for every $z \in K$ one has*

$$f(x) \leq f(z) \leq f(y).$$

In particular, if A is a relative compact in X then $f(A)$ is a bounded set.

Definition. A function $f : X \to Y$ is called uniformly continuous if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for every x, y satisfying $\rho(y, x) < \delta$ one has $\sigma(f(y), f(x)) \leq \varepsilon$.

Theorem 6.3 *Let X be a compact metric space and $f : X \to \mathbb{K}$ be a continuous function. Then f is uniformly continuous.*

Definition. Let X be a metric space. The space

$$\{g : X \to \mathbb{K} \mid g \text{ is continuous and bounded function}\}$$

with the metric

$$d_\infty(g, h) = \sup_{x\in X} |g(x) - h(x)|$$

is denoted by $C(X)$.

Exercise. Show that d_∞ is indeed a metric on that space.

Theorem 6.4 *The space $C(X)$ is complete.*

Proof: For simplicity we denote d_∞ just by d. First note that, by definition for every functions g, h and every $x \in X$ one has

$$|g(x) - h(x)| \le \sup_{z\in X} |g(z) - h(z)| \le d(g, h). \tag{6.1}$$

Let f_n be a fundamental sequence in $C(X)$, i.e. $d(f_n, f_m) \to 0$ as $n, m \to \infty$. By (6.1) it implies that for every (fixed) $x \in X$ one has $|f_n(x) - f_m(x)| \to 0$ as $n, m \to \infty$, i.e. $\{f_n(x)\}_n$ is fundamental in $\mathbb{K}$. Since $\mathbb{K}$ is complete we obtain that for every $x \in X$ the sequence $\{f_n(x)\}_n$ is convergent in $\mathbb{K}$. We denote the limit of that sequence by y_x and define the function $f : X \to \mathbb{K}$ by

$$f(x) = y_x = \lim_{n\to\infty} f_n(x). \tag{6.2}$$

To prove theorem it is enough to show that $f \in C(X)$ and that $d(f, f_n) \to 0$ (i.e. that f is limit of $\{f_n\}_n$ in $C(X)$. Note that by our construction we know only that f is the pointwise limit of $\{f_n\}_n$).

Fix $\varepsilon > 0$.

Since $\{f_n(x)\}_n$ is fundamental there exists $N_0 = N_0(\varepsilon)$ such that for every $n, m \ge N_0$ one has $d(f_n, f_m) < \varepsilon$, that is $|f_n(x) - f_m(x)| < \varepsilon$ for every $x \in X$.

Now, by (6.2), for every $x \in X$ there exists $N_1 = N_1(\varepsilon, x)$ such that for every $m \ge N_1$ one has $|f_n(x) - f_m(x)| < \varepsilon$.

It follows that for every $n \ge N_0$, for every $x \in X$, taking any $m > N_1$, one has

$$|f_n(x) - f(x)| \le |f_n(x) - f_m(x)| + |f_m(x) - f(x)| < \varepsilon + \varepsilon = 2\varepsilon.$$

Therefore for every $n \geq N_0$ one has

$$d(f_n, f) = \sup_{x \in X} |f_n(x) - f(x)| \leq 2\varepsilon. \tag{6.3}$$

Since f_{N_0} is continuous for every $x_0 \in X$ there exists $\delta = \delta(x_0, \varepsilon) > 0$ such that for every x satisfying $\rho(x, x_0) \leq \delta$ one has $|f_{N_0}(x) - f_{N_0}(x_0)| < \varepsilon$. Thus for every x satisfying $\rho(x, x_0) \leq \delta$ one has

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_{N_0}(x)| + |f_{N_0}(x) - f_{N_0}(x_0)| + |f_{N_0}(x_0) - f(x_0)| \\ &< 2\varepsilon + \varepsilon + 2\varepsilon = 5\varepsilon. \end{aligned} \tag{6.4}$$

Since f_{N_0} is bounded there exists $M = M(f(N_0)) > 0$ such that

$$\sup_x |f_{N_0}(x)| \leq M.$$

Therefore, by (6.3), for every $y \in X$ one has

$$\begin{aligned} |f(y)| &\leq |f(y) - f_{N_0}(y)| + |f_{N_0}(y)| \\ &\leq \sup_{x \in X} |f_{N_0}(x) - f(x)| + \sup_{x \in X} |f_{N_0}(x)| \leq 2\varepsilon + M_0. \end{aligned} \tag{6.5}$$

Since $\varepsilon > 0$ was arbitrary, (6.5) means that f is bounded, (6.4) means that

$$\forall \varepsilon > 0\, \forall x_0\, \exists \delta = \delta(x_0, \varepsilon) \text{ such that } |f(x) - f(x_0)| < 5\varepsilon \text{ whenever } \rho(x, x_0) \leq \delta,$$

i.e. f is continuous. Therefore $f \in C(X)$. Finally, (6.3) means

$$\forall \varepsilon > 0\ \exists N_0 = N_0(\varepsilon)\ \forall n \geq N_0 \quad \text{one has} \quad d(f_n, f) \leq 2\varepsilon,$$

which shows that f is the limit of $\{f_n\}_n$ in $C(X)$. It completes the proof. $\square$